

# MEASURING THE MONETARY VALUE OF LIFESAVING PROGRAMS

JAN PAUL ACTON\*

## INTRODUCTION

A multitude of public investment and regulatory decisions which have some effect on mortality and morbidity rates are made by legislatures, administrative agencies, and courts every year. Typically, as in the case of highway safety engineering, the choice which confronts the public decision maker is between reduced mortality rates, and hence longer life expectancy for some group, and more resources available for other purposes (*e.g.*, additional miles of highway construction or a reduction in taxes). A decision to require something other than the minimum technologically feasible mortality rate reflects in effect a judgment that mortality (or safety) is not to be given lexical priority in public decisions over all other commodities which money can buy—a judgment which is certainly reasonable and in accord with everyday decisions made by households. If mortality is not to be given lexical priority, some other standard or procedure is needed to determine which projects are worthwhile. In particular, a procedure is needed for measuring the benefits of such programs in units which can be readily compared with the costs.<sup>1</sup>

---

\* Economist, The Rand Corporation, Santa Monica, California. The author wishes to acknowledge with gratitude the comments of P. Cook, W. Manning, B. Mitchell, J. Newhouse, J. Vaupel, M. Weinstein, and A. Williams. The views are those of the author and do not necessarily reflect those of the Rand Corporation or any of its corporate sponsors.

1. Formal prospective evaluation of governmental programs, as discussed here, is a relatively young discipline. Water resource allocation has the longest history in the United States, having been charged since the 1930s to determine "if the benefits to whomsoever they accrue are in excess of costs." Flood Control Act of 1936, Pub. L. No. 74-738, 49 Stat. 1570 (codified at 33 U.S.C. § 701a (1970)). See Prest & Turvey, *Cost Benefit Analysis: A Survey*, in 3 *SURVEYS OF ECONOMIC THEORY* 155, 156 (1966). Most of these applications in water resources have been limited to economic benefits and costs, although considerations such as recreational values and their contribution have been added. See, *e.g.*, Weisbrod, *Income Distribution Effects in Benefit Cost Analysis*, in *PROBLEMS IN PUBLIC EXPENDITURE ANALYSIS* 177-209 (S. Chase ed. 1968).

A number of economists have reviewed various aspects of the program evaluation literature. Prest and Turvey, *supra*, have a good background regarding cost-benefit literature. P. STEINER, *PUBLIC EXPENDITURE* (1969), focuses on a number of issues in program budgeting for federal programs. A review of literature related to health evaluation, focusing on the evaluation of health technology is found in Klarman, *Application of Cost-Benefit Analysis to Health Systems Technology*, in *TECHNOLOGY AND HEALTH CARE SYSTEMS IN THE 1980s*, at 225 (DHEW Pub. No. HRA 74-3011, M. Collen ed. 1973). R. Thaler, *The Value of Saving a Life: A Market Estimate* (1974) (unpublished dissertation in Department of Economics, University of Rochester) reviews some historical attempts at valuation of lifesaving, and Zeckhauser, *Procedures for Valuing Lives*, 23 *PUB. POLICY* 420 (1975) provides a discussion of some recent applications. There are several essays on public expenditure in general. Dorfman and Chase have edited works which focus on particular prob-

In some constrained decision situations, the costs can be expressed in units of an identified commodity. A school board, for example, may be faced with the decision of how much of its budget to spend on school bus safety, knowing that every additional dollar spent on bus monitors and drivers' salaries will reduce the quality of education by a certain amount. The choice between safety and the quality of education is easily understood and could be assessed directly according to the preferences of the public as represented by the school board. More generally, money allocated to safety will be taken from a fungible source which has many alternative uses. In such cases, there is no good alternative to measuring the cost of safety in dollar terms, so that the evaluation of such a program will require the decision maker to place a dollar value on safety, at least in an implicit sense. Even in the school bus safety example, it is not appropriate to phrase the safety valuation question in terms of educational quality units if changing school taxes is a feasible option.

How is society to go about placing a dollar value on the health and safety effects of a public program? The method which is in accord with the theoretical postulates of welfare economics is to measure benefit as the sum of all affected individuals' willingness to pay for the proposed program.<sup>2</sup> One can imagine each household being informed of the potential effect of the proposed program on its members' own safety and the safety of all those they care about, and then sending a ballot to the appropriate agency which indicates the maximum amount they would be willing to pay to have the program enacted. Their response will reflect the risk aversion, their anxiety of dying from the particular cause which is to be modified by the program, their financial circumstances, and the objective reduction in risk to them and their friends. If the aggregate willingness to pay exceeds the costs of the program, then the program is worthwhile in the sense that everyone could be made better off by its adoption: It is possible, though probably not administratively practicable, to charge each beneficiary less than it is worth to him and still

---

lems of public expenditure evaluation. See MEASURING BENEFITS OF GOVERNMENT INVESTMENTS (R. Dorfman ed. 1965); PROBLEMS IN PUBLIC EXPENDITURE ANALYSIS (S. Chase ed. 1968). A set of essays on the Planning, Programming, Budgeting System (PPBS) experience by a number of practitioners and critics is found in R. HAVEMAN & J. MARGOLIS, PUBLIC EXPENDITURES AND POLICY ANALYSIS (1970). Some of the most extensive and successful applications of formal analysis have been in the defense area. Although they have tended to be cost-effective rather than cost-benefit analysis (*i.e.*, How can we best achieve a defense or tactical or strategic posture without asking how expensive a posture we should have?), some techniques developed there form the basis for analysis, especially regarding the general structuring of decision making under uncertainty and the quantification of uncertain outcomes. A good introduction to this systematic approach to analysis, with a description of a variety of techniques, is found in a collection of essays, SYSTEMS ANALYSIS AND POLICY PLANNING: APPLICATIONS IN DEFENSE (E. Quade & W. Boucher eds. 1968).

2. See generally Mishan, *Evaluation of Life and Limb: A Theoretical Approach*, 79 J. POL. ECON. 687 (1971). An interesting discussion of whose interests should be reflected in benefit valuation which considers the intergenerational problem is to be found in Dowie, *Valuing the Benefits of Health Improvement*, 9 AUSTRALIAN ECON. PAPERS 21 (1970).

cover the program costs. This "potential Pareto improvement" criterion is the formal theoretical justification for cost-benefit analysis, and it applies as well to evaluation of programs to reduce mortality or morbidity as to more traditional subjects like irrigation evaluation.<sup>3</sup>

This method, then, would define the benefit of a program which can be expected to save ten "statistical" lives out of a population of 100,000 as the total value the 100,000 members of this population place on having the probability of each individual's death reduced by one in ten thousand. An alternative method, and the one which is actually used in almost all evaluations of public health and safety programs, is to attempt to actually place a money value on the lives that the program would be expected to save if it were adopted. In the example above, the "benefit" of the program would be  $10V$ , where  $V$  represents the average "value of a human life." The method frequently used in practice for the heroic job of assessing  $V$  is to calculate the so-called "livelihood" measure<sup>4</sup>—the present value of lifetime earnings for a representative individual. The normative viewpoint which apparently motivates this approach is either that (1) people are properly thought of as the chattel of the state, and the loss of a life has a cost to the state comparable to the cost of a slave's death to his owner; or (2) the proper objective of public policy is to maximize gross national product.<sup>5</sup>

A third procedure for benefit valuation has not been employed in the past, but is potentially valuable. Since various public agencies and legislatures have been confronted with many decisions which in effect involve tradeoffs between dollars and mortality rates, there is considerable precedent for current decisions of a similar sort. Analyzing these precedents could help to increase the consistency of government decision making.

Before proceeding to discuss these basic approaches to measuring the benefit of safety enhancing programs in more detail, it is useful to indicate some of the seemingly related issues which, from a normative viewpoint, are in fact quite different. First, this article does not deal with the question of how much the government should spend to attempt to save the life of an identified individual (the coal miner trapped in a cave-in or the child in kidney failure) who is certain to die in the absence of government intervention. This is a very difficult issue because, among other things, there is a symbolic importance in

---

3. This criterion was originally proposed by both Kaldor, *Welfare Propositions of Economics and Interpersonal Comparisons of Utility*, 49 *ECON. J.* 549 (1939) and Hicks, *The Foundations of Welfare Economics*, 49 *ECON. J.* 696 (1939). A good recent discussion in the "valuing lives" context is Hirshleifer, *The Economic Approach to Risk-Benefit Analysis*, in *RISK-METHODOLOGY AND APPLICATIONS* (D. Okrent ed. 1975).

4. A term attributable to Schelling, *The Life You Save May be Your Own*, in *PROBLEMS IN PUBLIC EXPENDITURE ANALYSIS* 127 (S. Chase ed. 1968), as distinct from the lifesaving, or willingness-to-pay, approach.

5. See Mishan, *supra* note 2.

maintaining a public commitment to the preservation of life which, according to Calabresi and others, is properly viewed differently from the safety investment issue.<sup>6</sup> Second, this article does not attempt to determine the appropriate amount of compensation or punitive damages awarded to either the individual or his survivors for injury or death. While this is a related issue, in that court settlements in such cases may well influence the amount which private firms and households invest in safety, the relationship is complicated by equity considerations and a number of other factors, including the desire to establish correct incentives for people whose actions influence mortality rates.<sup>7</sup> Third, this article does not attempt to analyze the demand for life insurance, since this is determined by an individual's bequest motive and not by the value he places on his own safety.<sup>8</sup>

The remainder of this article considers each of the procedures for benefit valuation mentioned above, but in reverse order. A final section summarizes the principal arguments and makes several recommendations for policy analysts.

## I

### POLITICAL PRECEDENT

The logical first place to look for a source of standards for evaluating public programs which enhance health or safety is to the political process. If decisions regarding these programs tend to reflect a consistent set of values, then these values have a claim to political legitimacy and should be brought to light.

First, what does it mean for these decisions to be internally consistent? Investment and regulatory proposals differ in many dimensions, including the identity of the target population, the cause of death or disability which is to be curtailed, the nature and magnitude of the projected effect,<sup>9</sup> the various

---

6. G. CALABRESI, *THE COSTS OF ACCIDENTS: A LEGAL AND ECONOMIC ANALYSIS* (1970).

7. R. POSNER, *ECONOMIC ANALYSIS OF THE LAW* (1972) and McKean, *Products Liability: Implications of Some Changing Property Rights*, 84 Q.J. ECON. 611 (1970) have explored conditions under which economic efficiency is improved by assigning liability to one party (say the producer of a good) rather than permitting the market to supply (or fail to supply) products that provide reductions in risk. Although, in general, these liability solutions imposed to improve economic efficiency will understate the value of lifesaving or disability saving that would be inferred from a direct assessment of willingness to pay, they cannot be used as an unambiguous lower bound because of transactions costs and lack of perfect information, possible differences between the group determining the law and those engaged in the transaction, punitive elements to settlements, or differences between the group affected ex ante and the group being compensated ex post.

8. Eisner & Strotz, *Flight Insurance and the Theory of Choice*, 69 J. POL. ECON. 355 (1961).

9. Cohen, *Livelihood Benefits of Small Improvements in the Life Table*, in *HEALTH SERVICES RESEARCH* 82 (D. Maitland ed. 1975), reminds us that it is crucial to make clear the time course of the benefit for epidemiological as well as valational reasons. Frequently, analysts have in mind a program that offers a reduction in probability of death that is effective for one year at a time. Cohen points out that some program benefits may be more accurately characterized in a different

side effects, and the cost. To focus on the implicit valuations which such decisions place on improved mortality rates, two assumptions are useful: (1) Linearity: A program which reduces the probability of death by two in one thousand for each member of a specified group is worth twice as much as a program which causes only a one in one thousand reduction; and (2) Indifference to cause: The particular source of death which is to be curtailed by a program does not influence the program's value—all that counts is the number and perhaps the characteristics of lives saved. If these assumptions are accepted, then a consistent procedure for assessing the benefit of programs is to value each of them by the number of lives which it is predicted will be saved, multiplied by some number representing what is often called the average "value of life" for the program's target population.<sup>10</sup> Precedent decisions can be analyzed to ascertain whether they reflect a consistently applied set of life values.

For any number of reasons it comes as no surprise that public program choices do not reflect the type of consistency defined above. One study which examined a number of lifesaving programs found implicit values of life that ranged from a few thousand dollars (in highway safety design) to over a million dollars (in an ejection system for an air force bomber).<sup>11</sup> To some extent

---

manner, and that the alternative definition may make a large difference in the measured benefit. He defines a "curative" benefit as one that offers a person a one-time save (or reduction in probability of death) from a disease, regardless of the age at which it occurs, and then the person falls back into the normal risk pool. He defines a "preventive" benefit as one that eliminates a particular cause of death entirely. Cohen shows that substantial differences can arise in the measured total benefit when a curative or preventive benefit rather than a one-year exposure benefit is involved. In the case of kidney disease for American males, his calculations yield a total benefit about twenty-two times as large as that of J. HALLAN, B. HARRIS, & A. ALHADETT, *THE ECONOMIC COST OF KIDNEY DISEASE AND RELATED DISEASES OF THE URINARY SYSTEM* 8 (U.S. Public Health Service Pub. No. 1940, 1968).

10. It should be noted that while the "value of life" terminology is convenient and frequently encountered within the philosophical framework of the livelihood procedure, it is strictly accurate only because of the linearity assumption. If decision makers are non-linear with respect to livelihood saving (e.g., if they are not indifferent between (a) saving one person's life (and livelihood) with certainty and (b) saving one hundredth each of one hundred persons' lives and livelihoods), then one cannot even speak of the "value of a life" within the context of the livelihood measure. Within the context of willingness-to-pay measures, it is meaningless to speak of "the value of a life." In general, one can only refer to the expected value per life saved at a given initial risk of death and for a given reduction in risk. Suppose a given individual has an initial risk of death  $P$ , and is offered a chance to reduce it by  $\Delta P$ . If he will be willing to pay an amount,  $X$ , to reduce the risk, then we may refer to the value  $Y$  (which equals  $X/\Delta P$ ) as the expected value per life saved for this set of circumstances. (It can also be viewed as the amount that a large number of people similarly affected and with similar tastes would pay, on the average, for each life saved in their group.) In general (because of risk aversion and because one's budget constraint is affected by non-trivial changes in risk of death), people will not be willing to pay an amount  $2X$  for a reduction in risk of  $2\Delta P$ . Similarly, people whose initial risk is  $Q$  instead of  $P$ , will generally be willing to pay something other than  $X$  for the same  $\Delta P$ . We discuss some evidence about amounts people are willing to pay for different values of  $P$  and  $\Delta P$  in section III.

11. J. Carlson, *Valuation of Life Saving* (1963) (unpublished dissertation at Harvard University).

this variability may reflect deviations from one or both of the simplifying assumptions stated above. For example, a disproportionately higher and more expensive standard of safety for airplanes vis-à-vis highways may be justified by the argument that the threat of a crash seems to produce greater anxiety in air passengers than in auto passengers, even though the objective probabilities of death/mile are lower for the former group. In this vein one could also point to the disproportionate concern about death by shark bite or being murdered by a stranger.

Inevitably, however, much of the variability is the result of decentralized and varied decision-making processes, special political interests, and ignorance. Analyzing past decisions for precedents in defining the appropriate value of safety and health programs would be useful to the extent that it helped dispel this ignorance and yielded understanding of the implications of consistency for decisions concerning programs under current consideration.

Ultimately, the study of precedent decisions does not yield an absolute standard by which to measure benefits of potential programs. It does, however, offer a contingent standard which may be useful. If established program *X* is generally recognized as worthwhile, and proposed program *Y* offers a comparable increase in life expectancy/dollar expended, then there is a good argument for adopting program *Y*. In the absence of a consistent set of values generated by the political decision process, however, there remains a pressing need for benefit values calculated on the basis of more fundamental normative considerations. It is this need which, rightly or otherwise, is currently being filled by the "livelihood" procedure for life valuation.

## II

### LIVELIHOOD SAVING MEASURES OF VALUE

Livelihood saving is the most commonly used formal method for assessing the value of reducing mortality, and has been used as such for over fifty years.<sup>12</sup> This measure is based on the net present value of changes in the person's earnings stream.<sup>13</sup> By this criterion, if the expected livelihood savings

---

12. See, e.g., Crammond, *The Cost of the War*, 78 J. ROYAL STATISTICAL SOC'Y 361, ser. A (1915); Boag, *Human Capital and the Cost of the War*, 79 J. ROYAL STATISTICAL SOC'Y 7, ser. A (1916). For a review of some relevant literature, see L. DUBLINE & A. LOTKA, *THE MONEY VALUE OF MAN* (2d ed. 1946); Rice, *Estimating the Cost of Illness*, 57 AM. J. PUB. HEALTH 424 (1967). More recently, the livelihood saving approach has been used in a number of governmental evaluation studies. See, e.g., U.S. DEP'T OF HEALTH, EDUCATION, AND WELFARE, *DISEASE CONTROL PROGRAMS: SELECTED DISEASE CONTROL PROGRAMS* (1966); U.S. DEP'T OF HEALTH, EDUCATION, AND WELFARE, *HUMAN INVESTMENT PROGRAMS: SELECTED HUMAN INVESTMENT PROGRAMS* (1966). Reviews of its general application to other areas of analysis appear in Kiker, *The Historical Roots of the Concept of Human Capital*, 74 J. POL. ECON. 481 (1966); L. THUROW, *INVESTMENT IN HUMAN CAPITAL* (1970). One most extensively applied set of livelihood tables can be found in Rice & Cooper, *The Economic Value of Human Life*, 57 AM. J. PUB. HEALTH 1954 (1967).

13. That is, if the earnings in year *i* are  $E_i$ , the probability of surviving until year *i* is  $P_i$ , and

associated with a project exceed the costs of the project, it is worth undertaking;<sup>14</sup> otherwise the project is not worthwhile. Despite considerable discussion and use of livelihood saving measures in the literature, a clear statement of why it might be desirable to employ such a criterion for funding public programs does not appear. In particular, there is no reason to believe a priori that changes in earnings streams bear any direct relationship to what society values in health or safety program outputs.<sup>15</sup>

The livelihood saving approach may have received the attention it has because it is relatively easy to apply and gives the impression of providing an unambiguous numerical answer. It is easy because the analyst can consult a table to determine the livelihood at different ages, identified by sex, race, and education.<sup>16</sup> The impression of numerical precision is more apparent than real, however. A number of important assumptions underlie the tables, and unless the decision maker is conscious of their meaning, he may be unconsciously supporting a social judgment that he would reject if he faced it explicitly.

#### A. Intrinsic Shortcomings of Livelihood Approaches

The major objection to a livelihood evaluation is that it lacks a satisfactory normative justification. It is possible to infer from the way this approach is discussed in the literature that it is supposed to be justified by analogy to the economic procedure for valuing a machine or other piece of capital equipment. If a machine is accidentally destroyed, the resulting economic loss is equal to either (1) the cost of replacing the machine, or (2) the present value of the services which the machine would have provided if it had been saved

---

the discount (or interest) rate is  $r$ , then the livelihood of a person  $n$  years old is:

$$\sum_{i=n}^{\infty} \frac{P_i E_i}{(1+r)^{i-n}}$$

The choice of the discount rate,  $r$ , is not always unambiguous and may vary year by year.

14. That is, it passes the criterion of absolute merit. If resources are being rationed, there may be other uses of funds that generate an even greater net change in livelihood.

15. See Conley, *The Value of Human Life in the Demand for Safety*, 66 AM. ECON. REV. no. 1, at 45 (1976). Conley has recently argued that changes in expected present value of earnings provides a lower bound to individual willingness to pay for lifesaving programs. This conclusion is misleading, however, first because his analysis assumes that changes in life expectancy do not permit a reallocation of the consumption stream (that is, he assumed that the partial derivative of expected income with respect to a change in probability of survival is zero—which runs counter to his assumption of perfect annuity markets). Second, it has been pointed out that Conley's assertion rests on the sign of a set of terms in his first order conditions and that, although empirical evidence is lacking on either side, it is plausible that this set of terms would be negative for many people rather than unambiguously positive as Conley asserts.

16. Widely used sets of such tables are found in Rice & Cooper, *supra* note 12, and in Cooper & Brody, *1972 Lifetime Earnings by Age, Sex, Race, and Educational Level*, in U. S. DEP'T OF HEALTH, EDUCATION, AND WELFARE, RESEARCH AND STATISTICS NOTE (1975).

—whichever is less. If the market for such machines is competitive, then measures (1) and (2) are equal, and both valid. Furthermore, the value of the machine's services is equal to the implicit or explicit rental price of the machine. People can be viewed as embodying "human capital," the services of which are rented in the labor market or used in home "production" (house-cleaning, child care, etc.). The rental rate (wage rate) for labor services will, under some assumptions, reflect the value of such services in production. If one accepts the notion that the social value of a life is equal to the value of the labor services the person provides, then the present value of the person's expected earnings (including "implicit" earnings from home production) is the appropriate measure of this value.

People, however, are not machines. If one accepts the view that production is not an end in itself for people, but rather a necessary intermediate step which allows people to enjoy the fruits of production, then the "human capital" approach is clearly inappropriate. Increases in safety and life expectancy help to ensure the continuation of an individual's ability to enjoy the pleasures of his life and the continuation of the pleasure which his family and friends derive from their relationship with him, and it is the value of prolonging this enjoyment which should be assessed in measuring the benefit of public programs that affect safety. While this hedonistic view would not be appropriate in a slave society (at least from the owner's viewpoint) or in a society dedicated solely to increasing the gross national product, it seems entirely appropriate in an individualistic society where the government is viewed as serving the public rather than vice versa.<sup>17</sup>

The livelihood procedure might still be accepted in practice if it could be demonstrated that it provides a reasonable approximation to a measure which does have conceptual validity, even if only to our intuitive notions of what equitable policy requires. For some judgments at least, this type of justification is clearly lacking. It is, for example, an inescapable conclusion of this criterion that society should spend no money on programs that extend the lives of fatally ill children because the programs would produce no change in their future earnings. Furthermore, most persons would not agree that it is as

---

17. The logical extension of the viewpoint which seems to motivate the livelihood procedure is to argue that an individual's consumption should be deducted from his earnings in calculating the value of his life—that his value is equal to the present value of the surplus he generates (note again the analogy with the slave). One implication of this "net livelihood" procedure is that society is made better off by the death of those whose expected net present value is negative. This is true of retired people and those who are near retirement, some of whom receive disability and public assistance payments, some children, and so on. Dissatisfaction with the implied judgment that society should not expend any effort to extend the lives of such people has led researchers to use income without excluding consumption. See, e.g., R. FEIN, *THE ECONOMICS OF MENTAL ILLNESS: A REPORT TO THE STAFF DIRECTOR* (1958); Klarman, *supra* note 1; M. Feldstein, *Cost-Benefit Analysis and Health Planning in Developing Countries* (1970) (unpublished discussion paper on file at Harvard University).



important to save one worker earning \$10,000 per year as it is to save two workers with similar personal and family characteristics, but each earning \$5,000 per year. It is even more doubtful that most decision makers would want to save men and women in proportions that depend on their earnings, even if a homemaker's services are valued at the wages of a domestic worker rather than at zero. For instance, the livelihood saving calculation presented below shows that a white man in his fifties is valued more highly than a white woman in her twenties. If livelihood saving were used as the measure of value for government health programs, this means that programs would be approved that save fifty-five-year-old men rather than programs that save the same number of twenty-five-year-old women. It also indicates that it is worth about twice as much to save one twenty-five-year-old man as to save one twenty-five-year-old woman.

It is doubtful that these magnitudes reflect the rate at which most people would want public lifesaving and morbidity saving resources allocated. There is little direct evidence on this point about societal preferences, but what exists explicitly contradicts this implication of the livelihood approach. In Acton,<sup>18</sup> ninety-one persons were asked hypothetical questions about which person they would like to see saved if two seriously injured men arrived at an emergency ward and there were resources available to save only one of them.<sup>19</sup> The respondents had to choose between several different pairs of ages. Approximately one-third (thirty-one) of the respondents always chose to save the younger person; thirty-nine expressed a preference that was single-peaked in age (peaks generally occurred between twenty and thirty years of age as does the human capital curve); and eight were indifferent to all age pairs. (The remainder were multipeaked or inconsistent rankings.) Thus, somewhat less than half of the respondents expressed a desire to save lives identified by age that corresponds to the shape of the livelihood curve.

The livelihood measure assigns a higher value to men than to women at almost all ages, but this sample rejected such a ranking when asked to select a man or a woman of identified ages in the emergency-ward question above. The majority of persons (fifty-three) selected only on the basis of age and matched the same ranking they had expressed when selecting between two men. Nine respondents always selected the man over the woman, and nine always selected the woman over the man. In one question, the respondents were asked to choose between a thirty-year-old man and a thirty-year-old

---

18. J. ACTON, *EVALUATING PUBLIC PROGRAMS TO SAVE LIVES: THE CASE OF HEART ATTACKS* (1973).

19. Thirty-six of these respondents were selected at random from three communities in Boston (half men and half women); nineteen were men in a trade union program, and thirty-six were in an advanced management program at the Harvard Business School. For a description of these samples, see *id.* at 83-85.

woman. Thirty-seven chose the man, forty-three chose the woman, and eleven expressed indifference.

I am not aware of any other systematic empirical evidence about people's preferences for saving lives identified by age or by sex. However, this empirical evidence, along with casual observation of attitudes for public programs, suggests that a majority of people would at least reject the relative value of saving men and women that is implied by the simple livelihood method. In the provision of public services, where objectives may include allowance for factors such as income redistribution, and externalities such as the number of dependents that will be orphaned, the social evaluation may even vary inversely within measures of livelihood involved!

Even if one were satisfied that the livelihood procedure formed a conceptually sound basis for public program evaluation, an important practical issue remains to be resolved: Market earnings in some cases do not equal the productivity of an individual's labor.

#### B. The Issue of Earnings vs. Productivity

A person's earnings may differ significantly from his productivity for a number of reasons. For instance, workers in a strong union may earn considerably more than workers doing identical, nonunionized work. Some groups may face earnings discrimination because of race, ethnicity, or sex. Some people (*e.g.*, those with job seniority) may be receiving an income substantially above their productivity. The livelihood measure is blind to these distortions. It merely says to add up the earnings of people who may be affected by different programs, and select the ones that save the most earnings. Since diseases typically do not affect different racial, sexual, or socioeconomic groups uniformly, a criterion that depends on earning differences among these groups will necessarily slant public programs in particular directions. If some diseases are found more often in people with higher earnings, the rule says to devote attention and resources to these diseases.

The undesirable nature of this criterion is brought home acutely when one considers the implications for the treatment of women (although it applies in less extreme form to any case where wages do not reflect productivity). The national product accounts do not include the homemaker services of women if they are not purchased; but to exclude them from a measure of project benefit will seriously undervalue programs that affect women. The most common procedure is to value homemaker services as the full time earnings of a domestic worker.<sup>20</sup> Various arbitrary weighting rules have also been used.<sup>21</sup>

Using the earnings of a domestic servant is only partly satisfactory, how-

---

20. Compare Weisbrod, *The Valuation of Human Capital*, 69 J. POL. ECON. 425 (1961), with Klarman, *Syphilis Control Programs*, in R. Dorfman, *supra* note 1, at 367, and with Rice, *supra* note 12.

21. See, *e.g.*, M. Feldstein, *supra* note 17.

ever. In the first place, the homemaker may be providing quantity or quality of services that are not available in the market. For instance, when one observes a woman with advanced education who could take a job paying two or three times a domestic servant's income, she may be staying home to raise her small child because she feels the first few years are important and because she does not feel she could hire such high-quality nurturing for her child. Under these circumstances, using the domestic servant's earnings will understate the value of this woman's home activities, as she sees them. In such circumstances, one could argue that her services at home should be valued at least as highly as the highest salary the woman could earn.<sup>22</sup> However, one probably does not really wish to adopt the implications of such reasoning. After all, many people accept jobs at a salary less than the maximum they could command in the market. They may do this in order to have better working conditions or in order to pursue a particular type of work. In the extreme, the implication of this foregone opportunity argument is that one should value everyone's services—men's and women's—at the highest possible wage they could earn. Ignoring the readjustment this would cause in the general wage scale, such a recalculation would raise the implicit earnings of society considerably.

A second objection to the standard treatment of home production is that it is asymmetric with respect to sex. After all, women are not the only workers around the home. Morgan<sup>23</sup> and Walker and Gauger<sup>24</sup> surveyed people about the hours they spend working around the house. They found that men spend between one-eighth and one-third as much time as do women, depending on the employment status of the woman, and the ages and family sizes involved.<sup>25</sup> If one were to impute a value to individuals for their home production then it seems appropriate to add an element to the man's livelihood calculation.

The third objection to the standard treatment of home production lies in the treatment of older women, especially those over sixty-five years of age.

---

22. For instance, we could examine the earnings of women with similar education and training who are employed full time in the market and impute those earnings to the women who stay home. See R. POSNER, *supra* note 7, at 79-80.

23. J. MORGAN, I. SIRAGELDIN, & N. BAERWALDT, *PRODUCTIVE AMERICANS: A SURVEY OF HOW INDIVIDUALS CONTRIBUTE TO ECONOMIC PROGRESS* (University of Michigan Survey Research Center Monograph No. 43, 1966).

24. K. WALKER & W. GAUGER, *THE DOLLAR VALUE OF HOUSEHOLD WORK* (New York State College of Human Ecology Information Bull. No. 60, 1973).

25. Rice & Cooper, *supra* note 12, assumed that all nonemployed women contributed a full share to home production and assigned the full time earnings of a domestic worker to those women, about \$2767 per year in 1964. They assigned no other value for household production to others. This implies, among other things, that it is frequently better to save women who do not work than it is to save women who work part time. In Cooper & Brody, *supra* note 16, the value of housework measured by K. WALKER & W. GAUGER, *supra* note 24, was used, but no adjustment is made for men or for changed productivity after age sixty-five.

Rice and Cooper<sup>26</sup> attributed a full domestic worker's income to nonemployed women over sixty-five, causing their livelihood to exceed significantly that of a man over sixty-five. One could speculate that women over sixty-five start to slow down in their household activities, but it is difficult to find data. Walker and Gauger<sup>27</sup> did not survey older women. I analyzed the results of the Productive Americans Survey.<sup>28</sup> The number of observations is relatively small in the over sixty-five age group, but there appears to be a downturn in average number of hours worked at home by women and an increase in the hours worked by men. Women's hours declined about 19 per cent in the over sixty-five age group and men's hours increased about 17 per cent. This leaves women over sixty-five reporting about thirty-five hours of housework per week and men reporting about six and one-half hours. These figures may represent an overstatement of true contribution if productivity falls significantly in this age group. Furthermore, there may be some reporting error if the respondents have little else to do and therefore claim that most of their time goes to housekeeping.

Since there are no compelling theoretical arguments for one rule over another in accounting for household production, livelihood tables can be generated under a variety of assumptions about the value of women's and men's contributions.<sup>29</sup> These calculations show significant variation in the livelihood, especially in the upper ranges, depending on the assumptions employed. For illustrations, Figures 1 and 2 plot the livelihood at different ages for the four-way breakdown of sex and race under two of the assumptions possible for treating home production. The assumptions behind the calculations are discussed in more detail in Acton,<sup>30</sup> but, briefly, Figure 1 (Assumption 1-1) assigns a value of \$4800 for the domestic work of nonworking women.<sup>31</sup> Figure 2 (Assumption 3-3) assigns a variable amount to women's homemaker function (depending on their employment status) and a uniform amount to men. After sixty-four years of age, women's contribution is reduced 19 per cent to reflect a drop in household activities, and men's is increased 17 per cent. A four per cent net discount rate is used for both figures.

---

26. Rice & Cooper, *supra* note 12.

27. K. WALKER & W. GAUGER, *supra* note 24.

28. J. MORGAN, I. SIRAGELDIN, & N. BAERWALDT, *supra* note 23.

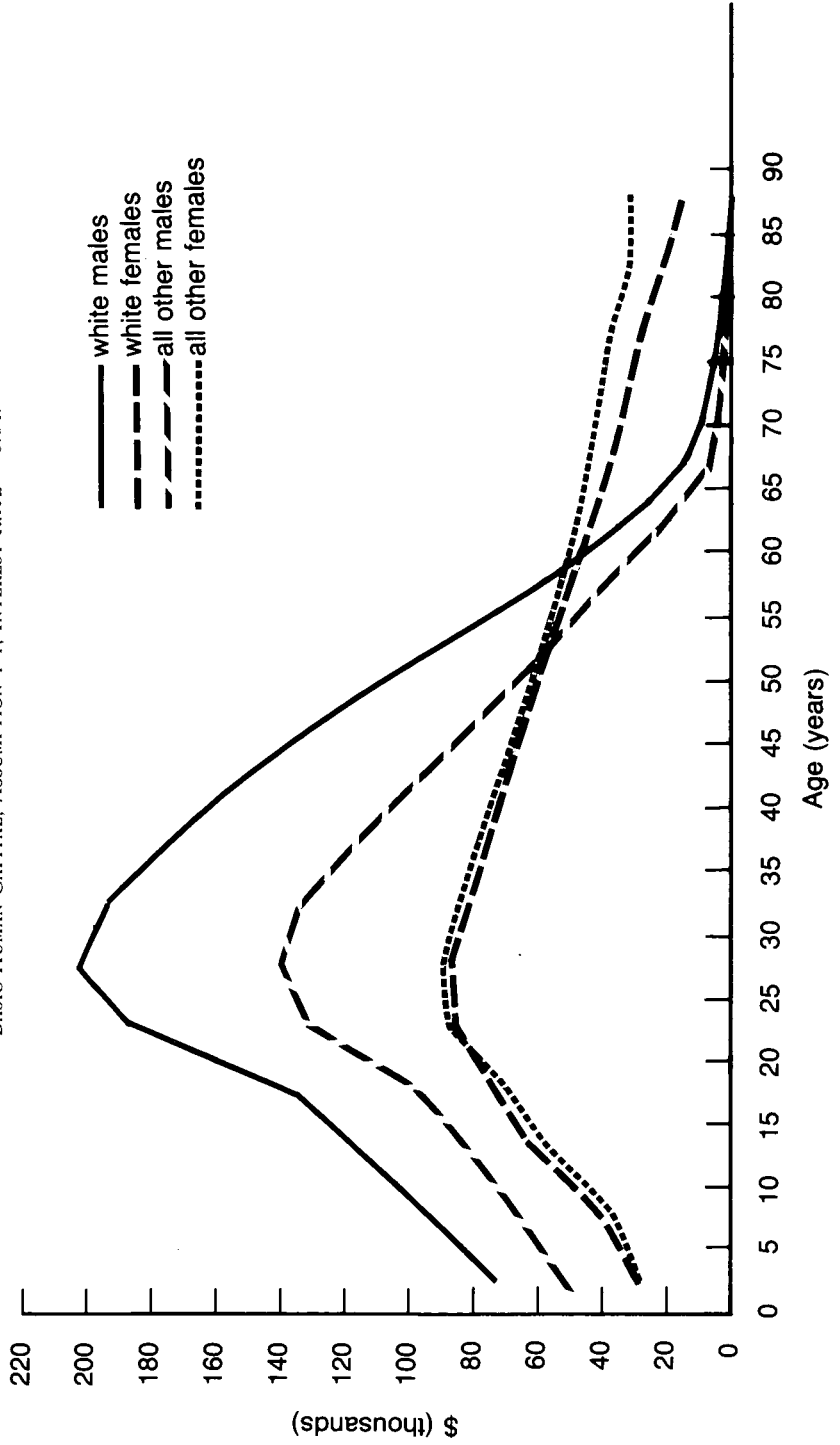
29. J. ACTON, MEASURING THE SOCIAL IMPACT OF HEART AND CIRCULATORY DISEASE PROGRAMS: PRELIMINARY FRAMEWORK AND ESTIMATES (1975).

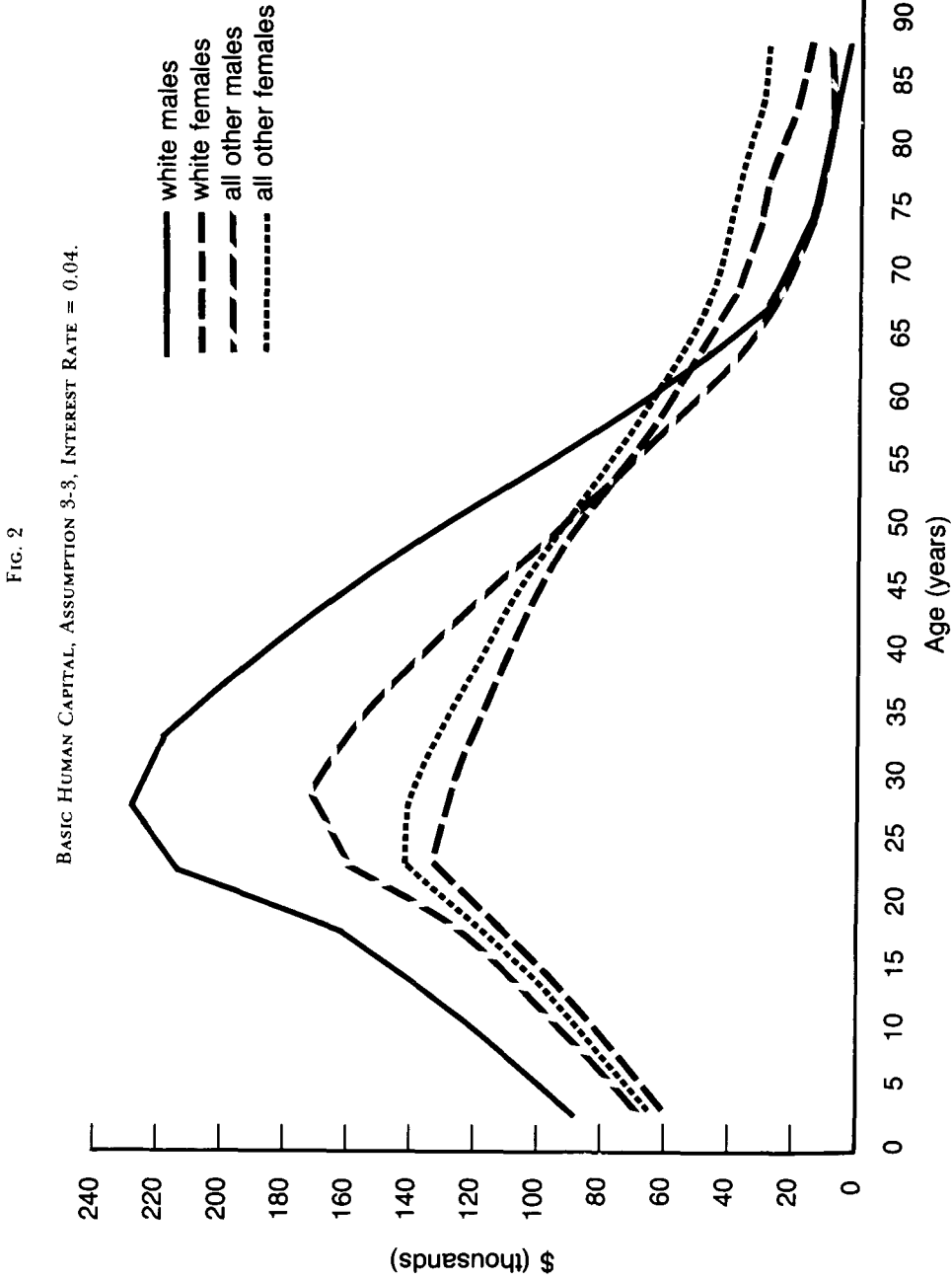
30. *Id.* § IV.

31. After this work was completed, Dorothy Rice (personal communication) informed me that the domestic worker's earnings for 1972 were about \$4000. Resources did not permit recalculation of all the human capital tables to adjust for this fact, but one should note that it does not change the character of the methodological and empirical findings. If recalculated, the differential between men and women would increase during the working years and narrow somewhat over sixty-five years of age. The average amount for willingness-to-pay measure would increase further over the human capital amount.

FIG. 1

BASIC HUMAN CAPITAL, ASSUMPTION 1-1, INTEREST RATE = 0.04.





I do not intend to focus on the nature of livelihood at different points in life or to concentrate on differences among races or between sexes (although they are already quite substantial). These plots, however, serve to emphasize the substantial variability which results from alternative assumptions about the valuation of household activities and the substantial impact this has on the relative and absolute amount assigned to women by this criterion. The effect of these alternative assumptions is significant at all ages, but is especially noteworthy in the over sixty-five age range where a substantial amount of mortality and morbidity is involved from such prominent ailments as heart and circulatory diseases and cancer.

The plots in Figures 1 and 2 show a close similarity between the livelihood for white females (WF) and all other females (AOF). This stems from the relatively low work rates of women, combined with the assumption that all nonworking women are assigned the same value of household services regardless of race. The differences between white males (WM) and all other males (AOM) is about the same under the two assumptions and measures about \$60,000 higher for white men in their late twenties than nonwhite men of the same age. The difference between sexes is dramatic—with the livelihood of white males at its peak equal to about two and one-half times the level of white females at its peak under Assumption 1-1. When the household production of working men and women is given an imputed value (Assumption 3-3), the differences between the sexes narrow considerably. At its peak, white men's livelihood is only 1.7 times that of white women. The male:female ratio is even closer for nonwhites.

The other major effect of the different assumptions comes in the cross-over between male and female livelihood in the upper age brackets. Under Assumption 1-1, female livelihood crosses male between fifty and sixty years of age due to the lower life expectancy of men and the fact that women are assigned a value of household production while the generally retired men are not. Consequently, over sixty-five years of age, male livelihood falls to extremely low levels, while female livelihood remains between \$20,000 and \$40,000. Under Assumption 3-3, when a greater value is assigned to household production for men and for working women, the reversal for white men's and women's livelihood is postponed to the early sixties, and the livelihood of men is higher than before in both relative and absolute terms. The reversal for nonwhites is pushed to a lower age, but the difference at all ages is narrowed considerably.

### III

#### THE WILLINGNESS-TO-PAY MEASURE OF VALUE

A fundamental assumption of the willingness-to-pay procedure is that individuals' preferences should count; that citizens can and should play a role

in policymaking for governmental services that affect them directly. Their health, their friends, their taxes, their pain and suffering, and their welfare are at stake. Understandably, they have an interest in the public activities that may be undertaken. Individuals are the ultimate recipients of the impact of programs.

Political justifications for using individual preferences go back at least to the seventeenth century and include the desire for no taxation without representation. Economic arguments for using an individual's preferences date from the nineteenth century and include the utilitarian principles of Bentham. Dupuit,<sup>32</sup> a French engineer, argued that the nature and amount of public transportation facilities should be determined by what the potential users would be willing to pay for using it. Most contemporary economists who study public policy evaluation agree that an approach based on individual values is correct in principle.<sup>33</sup>

The "potential Pareto improvement" standard which justifies the willingness-to-pay procedure has been questioned because it makes the estimated dollar benefit of a program dependent on the income distribution. This dependence has been criticized either because (1) it is felt that the income distribution is inequitable and hence not a just basis of public program evaluation; or (2) it is felt that whether or not the income distribution is equitable it is simply not an appropriate basis for determining the production and distribution of certain goods (possibly including adequate health care and safety) which are, like the vote, properly considered noncontingent privileges of membership in society.<sup>34</sup> The problem which has not been solved by critics is to devise an alternative benefit measure which satisfies such objections. The livelihood measure is even more directly tied to income distribution (*viz.*, by definition) than is the willingness-to-pay measure, and it is not impossible that precedent political decisions were influenced by the economic power of various interest groups.

The main practical problem with the willingness-to-pay procedure for benefit estimation is that developing accurate assessments of an individual's willingness to pay is difficult and expensive, and the validity of published attempts to apply various estimation techniques has not been demonstrated.

---

32. Dupuit, *On the Measurement of the Utility of Public Works*, in READINGS IN WELFARE ECONOMICS 255 (K. Arrow & T. Scitovsky eds. 1969).

33. M. FELDMAN, M. PIOT, & T. SUNDARESON, RESOURCE ALLOCATION MODEL FOR PUBLIC HEALTH PLANNING: A CASE STUDY OF TUBERCULOSIS CONTROL (1973); Bohm, *An Approach to the Problem of Estimating Demand for Public Goods*, 73 SWEDISH J. ECON. 55 (1971); Lave & Weber, *A Benefit-Cost Analysis of Auto Safety Features*, 2 APPLIED ECON. 265 (1970); Mishan, *supra* note 2; Samuelson, *Diagrammatic Exposition of a Theory of Public Expenditure*, 37 REV. ECON. & STATISTICS 350 (1955); Samuelson, *The Pure Theory of Public Expenditure*, 36 REV. ECON. & STATISTICS 387 (1954); Zeckhauser, *supra* note 1.

34. See A. OKUN, EQUALITY AND EFFICIENCY: THE BIG TRADEOFF (1975); Tobin, *On Limiting the Domain of Inequality*, 13 J. LAW & ECON. 263 (1970).



Furthermore, the extent to which estimates of a particular population group's willingness to pay for a particular safety-enhancing project can be applied to other groups and other types of projects is unknown.

The two principal methods for measuring the values a household would place on a prospective public project are: (1) Inferences of how much the household values mortality reduction based on observations of the implicit value the household places on safety and health in making private consumption and job-selection decisions; and (2) Survey questionnaires which ask household heads to state their willingness to pay for the program benefit that is under consideration.

#### A. Implicit Values

One can, in principle, infer the values individuals attach to mortality- and morbidity-reduction in the same manner as was proposed for governmental actions (section I above). Such a revealed preference approach is followed with most market-produced goods that have few externalities.<sup>35</sup> One need not go into a detailed survey of relative preferences for, say, apples and oranges. People reveal the preferences they attach by their market behavior. This is the method one would like to use to measure individuals' true preferences for the programs. It presents the strongest claim to validity because the people have to back up their preferences with action, and they do it in the context of other everyday decisions for spending money.<sup>36</sup> These choices may include the purchase of safety devices (for example, seat belts), a marginal expenditure on health items (perhaps a doctor's examination and some antibiotics for an infection), or the premium demanded for accepting an elevated risk (for instance, higher wages for extrahazardous employment).

Recent studies have provided measures of implicit willingness to pay for lifesaving. Thaler,<sup>37</sup> Thaler and Rosen,<sup>38</sup> and Smith<sup>39</sup> examine the higher wages paid in occupations with above-average risk of death for evidence about the implicit value of lifesaving. Usher<sup>40</sup> employs a life-cycle model of utility

---

35. That is, effects that extend beyond the principal economic agents. A good example of externalities is the pollution that may be generated in the production of some goods. Neither the manufacturer nor the consumer of the good pays for the smoke (at least until recently), although a number of people experience the effects, would like to see them reduced, and would be willing to pay to have them reduced.

36. Dreze, in particular, has argued the merits of using this procedure. See Dreze, *L'Utilité Sociale d'une Vie Humaine*, 23 REVUE FRANCAISE D'INFORMATIQUE ET DE RECHERCHE OPERATIONELLE 93 (1962).

37. Thaler, *supra* note 1.

38. Thaler & Rosen, *The Value of Saving a Life: Evidence From the Labor Market*, in HOUSEHOLD PRODUCTION AND CONSUMPTION 265 (M. Terleckj ed. 1975).

39. R. SMITH, COMPENSATING WAGE DIFFERENTIALS AND HAZARDOUS WORK (1973).

40. Usher, *An Imputation to the Measure of Economic Growth for Changes in Life Expectancy*, in THE MEASUREMENT OF ECONOMIC AND SOCIAL PERFORMANCE 193 (M. Moss ed. 1973).

maximization and infers the tradeoff between consumption and probability of survival from a time series of the national income accounts and mortality statistics. Both approaches have the potential of overcoming some reservations about the survey-based willingness-to-pay approach, because they examine behavior revealed through market activity and therefore have stronger claims to validity and stability than existent survey results.

Since the two Thaler studies and the Smith study rest on market wages, they have some drawbacks in common with the livelihood saving approach. First, the measure requires that the person be working to determine a value. Therefore, it is difficult to ascertain the appropriate value for housewives, children, retired persons, and others who are not paid for their work. A second criticism relates to the representativeness of this group observed in riskier occupations. Presumably, those who are least risk-averse will enter a given occupation before those who are more risk-averse, all other things the same. Consequently, lower risk premiums will be paid to those who select an occupation than to a randomly chosen individual who was subjected to the same level of risk, and these measures will be a lower bound on "society's value." Third, the extra pay is compensation for assuming an above-average risk, and for that reason may not provide an appropriate measure of value for programs which are designed to reduce risk. The compensation which a risk-averse person would require to accept a  $\Delta p$  increase in the probability of his own death is greater than the amount he would be willing to pay for a  $\Delta p$  reduction in this probability, although the amounts will be close to one another for small  $\Delta p$ . Fourth, the wage-premium observed will not necessarily reflect the externalities (to family and/or society) associated with a person's death, although the externalities will be better captured with this measure than with the livelihood saving approach if the employee includes his family in the job-choice decision and requires that the wage differential be adequate to compensate them for his increased risk as well. Fifth, it is difficult to identify what portions of differences in compensation correspond to the additional risk of death or injury or to other working conditions. Sixth, although it is not a general phenomenon, there may be some occupations in which the participants receive some utility from the risk, and therefore the compensation is inadequate for a normal person. Being a stock car racer or being a test pilot may be extreme examples, but this consideration may be reflected to some degree in a number of occupations, some of which are included in Thaler's calculations. Finally, at the conceptual level, one does not know for certain what risks of death or injury the individual assumed were in force when he accepted the wage offer. Given the difficulty Thaler seems to have had in getting good data on death rates by occupation, the amount of uncertainty a given individual faces about the risk at a particular job site may be substantial.

On the empirical side, Thaler found significant variation in implicit valua-

tion depending on the data source used. With one data file, he inferred a value of between \$176 thousand and \$260 thousand per expected life saved (for a reduction in probability of 0.001 per year), which is remarkably close to the peak human capital value observed for young men and to the explicit willingness to pay obtained by Acton<sup>41</sup> in his survey for a reduction of 0.001 in heart attack death rate. On the other hand, the value implicit in the Bureau of Labor Statistics injury data was over \$2.6 million per expected life. Furthermore, Thaler's regression results with the Bureau of Labor Statistics data yield an incorrect sign for the coefficient of risk of injury. The regression with the first data file did not include a variable for risk of injury, so his results are subject to omitted variable bias, and the differences between the first and second estimates are even more extreme than they appear.<sup>42</sup>

Usher's study is an imaginative use of the (Canadian) national income accounts to infer a tradeoff between consumption over a life cycle and resources devoted to death reduction. He makes utility solely a function of consumption in each time period (which is equal in all time periods) as well as the probability of surviving, and employs strong assumptions about the form of the utility function to make his estimates. Given the strong assumption about functional form, the potentially severe aggregation bias from using such highly aggregated data to infer a utility function for individuals, and the absence of an indication of the level of statistical significance, we may wish to place most emphasis on the qualitative findings. Usher's model implies that the value per expected life saved is greatest at a very young age (it peaks around age two for plausible values of his parameters) and decreases through increasing age. Its value in the age sample twenty to thirty is very similar to the human capital values reported for white males by Rice and Cooper.<sup>43</sup> Since utility is a function solely of consumption (not earnings) and since he assumes that every one consumes the same amount in each year of life, there is no difference between the value assigned to men and women in his model.

### B. Explicit Statements of Individuals

The survey approach<sup>44</sup> permits measurement of the entity which is directly appropriate for evaluating a proposed public project: the maximum amount each affected household would be willing to pay to have the project adopted. In theory this procedure requires no assumptions about individual preferences (*e.g.*, linearity, indifference to cause, absence of externalities)

---

41. J. Acton, *supra* note 18.

42. That is, risk of injury is probably positively correlated with risk of death. Omission of the first variable will bias the coefficient of the second variable away from zero, causing his estimates with the first data file to be too high.

43. Rice & Cooper, *supra* note 12.

44. Advocates of this approach include Schelling, *supra* note 4; V. TAYLOR, *HOW MUCH IS GOOD HEALTH WORTH?* (1969); J. ACTON, *supra* note 18.

which other techniques require. Since the expense of conducting a special survey for every proposed project would be prohibitive, however, in practice one would want to generalize from the results of one survey in order to assess other project proposals; such generalizations will of course require some assumptions on preferences.<sup>45</sup>

While willingness-to-pay surveys have been conducted successfully in recreation program evaluation,<sup>46</sup> the only published survey I have found of willingness to pay for health programs is contained in Acton,<sup>47</sup> and that survey deals only with programs that reduce chances of sudden accidental death or heart attack death. It sought preliminary evidence on the feasibility of applying willingness-to-pay responses to actual program evaluation and addressed several questions:

1. Can questions be formulated that in principle get at willingness to pay?
2. Do people seem willing to answer and are they relatively comfortable in answering such questions?
3. Are the responses people make subject to a rational interpretation?
4. What seems to be the major factors influencing stated willingness to pay?

---

45. Recently, a number of researchers have considered the nature of the utility function that may underlie an individual's willingness to pay for lifesaving. H. RAIFFA, *PREFERENCES FOR MULTI-ATTRIBUTED ALTERNATIVES* (1969) has shown under very general assumptions that a self-interested person, living alone (with no heir and a prepaid funeral), should pay more for a given reduction in probability of death if he is at a greater overall risk of death. J. Pliskin, M. Weinstein, & R. Shepard, *Utility Functions for Life Years and Health Status* (Oct. 1974) (unpublished paper on file at Harvard School of Public Health); M. Weinstein, R. Shepard, & J. Pliskin, *Decision-Theoretic Approaches to Valuing a Year of Life* (1975) (unpublished paper on file at Harvard School of Public Health) consider the valuing of life-years as a problem in multi-attributed utility theory, where the joint or conditional nature of the "good" being offered makes a difference to the inferred value. Cook & Graham, *The Demand for Insurance and Protection: The Case of Irreplaceable Commodities*, 41 Q.J. ECON. 143 (1977) explore the relationship between willingness to pay to avoid a loss and the compensation required to make a person as well off after a loss. Jones-Lee, *Valuation of a Reduction in Probability of Death by Road Accident*, 3 J. TRANSP. ECON. & POLICY 37 (1969) provides an analysis of the compensating variation required for various changes in the probability of death or injury. Usher, *supra* note 40, and Conley, *supra* note 15, formulate the issue as a life-cycle model in which the individual is assumed to try to maximize his expected lifetime utility, which depends directly on his consumption in each time period. Actual application is rare, however, since most writers have stopped with a theoretical treatment or have chosen an admittedly inferior technique for actual measurement.

46. Knetsch & Davis, *Comparisons of Methods for Recreation Evaluation*, in *ECONOMICS OF THE ENVIRONMENT* (R. Dorfman & N. Dorfman eds. 1972).

47. J. ACTON, *supra* note 18. Related work includes the survey of willingness to pay for selected disease entities conducted by M. Palmatier, *Willingness to Pay for Health Services* (Jan. 18, 1969) (unpublished paper in Department of Economics, University of Southern California). A prototype survey for determining individual tradeoffs among attributes of disease reduction programs was developed by E. KEELER, *MODELS OF DISEASE COSTS AND THEIR USE IN MEDICAL RESEARCH RESOURCE ALLOCATIONS* (1970); Berg, *Establishing the Values of Various Conditions of Life for a Health Status Index*, in *HEALTH STATUS INDEXES* (R. Berg ed. 1973); Torrance, Sackett & Thomas, *Utility Maximization Model for Program Evaluation: A Demonstration Application*, in *id.*, have some imputed values for medical risk-taking based on the responses of physicians in their role as proxy decision maker for patients.

In total, approximately 125 persons were questioned about their willingness to pay for heart attack mortality reduction.<sup>48</sup> People were posed four types of questions:

1. Age choice questions: Which of two seriously injured would you like to see saved in an emergency? Those results were discussed above in the critique of livelihood-saving measures.
2. Lives in the community: How much would you be willing to pay to have a heart attack ambulance that is expected to save  $X$  lives per year of the ten thousand people living around you?
3. Advice willingness to pay: Suppose your neighbor has just been told his risk of heart attack is  $Y$  per year, and his chances of dying if he has a heart attack are  $Z$ . How much do you think he should be willing to pay per year for a heart attack program that would reduce his chances of dying to  $Z^*$ ?
4. Own willingness to pay: Suppose your doctor tells you your chances of a heart attack are  $Y$  per year, and your chances of death, given the heart attack, are  $Z$ . How much are you willing to pay per year for a heart attack program that can reduce your chances of dying to  $Z^*$ ?

Each respondent answered twenty-six questions of type (1), two questions of type (2), and four questions each of types (3) and (4).

The results showed that it is possible to pose questions that get at the underlying issues of willingness to pay. Furthermore, people were willing to complete the interview and seemed relatively comfortable and responsive in doing so (the refusal and breakoff rates were negligible). The question of rational interpretation of the responses was not clearly resolved in a single survey of this size. Responses varied significantly from one individual to the next (only part of this could be explained as sampling variance due to sample size). High variation *per se* is neither unexpected nor undesirable for these types of questions. One expects preferences and attitudes to vary from one individual to the next, even for identical expected benefits offered to individuals who appear to be similar in the socioeconomic and demographic profiles. Nevertheless, the responses of most persons could be given a rational interpretation, and predicted effects were found for important explanatory variables such as income, wealth, age, and sex. The empirical results are discussed in detail in Acton.<sup>49</sup> Briefly, the principal statistically significant findings were that willingness-to-pay responses increase with increasing probability of death

---

48. Part of the sample was a representative community sample in the Boston area, and part was a sample of young and middle-aged men in a business school program. A variety of questionnaire forms were used so it is not possible to report empirical results for the full sample to identical questions. The questionnaire for these surveys is contained in J. ACTON, *supra* note 18 app.

49. *Id.* at 92-105.

and with greater reductions that are offered—but not in a linear fashion.<sup>50</sup> Second, willingness-to-pay responses are greater the more concretely and immediately the hypothetical program is related to the individual.<sup>51</sup>

If such willingness-to-pay responses were to be used routinely for program evaluation, one would wish to conduct a survey of a greater number of respondents (appropriately selected for statistical representativeness) where the questions included several different probabilities of mortality and morbidity and several different reductions in the values of each health consequence. If it appeared conceptually or empirically desirable, separate sets of questions for major categories of diseases or risks should be prepared (for instance, heart diseases, cancer, accidents, and so forth). If satisfactory, statistically significant willingness-to-pay relationships were found, then it would probably be most efficient to use the results of multivariate regression equations to estimate the aggregate willingness to pay associated with a particular program, taking account of the socioeconomic and demographic characteristics of the population affected and the anticipated changes in probabilities.

A number of issues are still left open in the feasibility of a survey-based method for eliciting value. These include the validity of the responses, their stability and replicability, problems with understanding and processing the information in these hypothetical situations, and strategic behavior in responding.

The validity of responses to willingness-to-pay questions has not been examined empirically. Indeed, it is not clear that the validity can ever be firmly established. A rigorous test of validity might be to survey a group of people and then come back and actually market the goods that had been described (say a heart attack ambulance) or raise their taxes in accordance with responses. Some people might refuse to act in accordance with their previous responses because of intervening factors which may be difficult to control against and which the respondent cannot even articulate.<sup>52</sup>

The stability and replicability of these preliminary results have not been demonstrated. Further empirical work is clearly needed to see if the same people respond with a reasonably stable set of preferences when resurveyed at a later date. Furthermore, it is still unknown if the results can be replicated in other geographic areas with different socio-economic and ethnic samples.

Several competing objectives exist in asking questions that are both realis-

---

50. This finding is further evidence that individual preferences do not follow the implications of a livelihood-saving measure, which is strictly proportional to income. We can infer both risk aversion and an upper limit on willingness to pay for a given mechanism of death reduction from these data.

51. That is, the responses to question types (2) were generally less than the responses to types (3), which were generally less than responses to types (4).

52. For instance, after he thinks over what it might be like to be confined to a bed for a long period of time, his willingness to pay to avoid such disability might change.

tic and yet understandable for the respondents. Since many of the situations posed to people are hypothetical (either the disease state or the consequences of the programs), the individual's comprehension of the situation is uncertain. For instance, although heart disease accounts for about one-third of all deaths per year, the realistic chance a person has of dying from a heart attack is less than one per one hundred per year for the majority of adults. Analysts are, as yet, uncertain about how well people understand and process such numbers.

Similarly, it is not necessarily known how well people understand the nature of certain disability states or recoveries. The operationally relevant point, however, is whether they understand the situation well enough during an interview that their preferences do not change significantly if a decision is made to inaugurate the program. The most direct way to test this assumption is to examine the stability of responses over time.

A fourth unresolved issue in willingness-to-pay elicitation is whether people will engage in strategic behavior when they respond. Lindahl<sup>53</sup> observed that people may have an incentive to underrepresent their true valuation if their taxes depend on their stated value. Acton<sup>54</sup> and Bohm<sup>55</sup> observed that the opposite case may also exist if people think the decision whether or not to have the program is based on aggregate value, but the cost-sharing rule is determined by a different standard. Under these circumstances, if a person feels he will be called on to bear a small proportion of the costs for a project he wants, he should overrepresent his willingness to pay for it. Dreze and Poussin<sup>56</sup> have shown that, under some circumstances, people will have the correct incentives to reveal their true preferences for public goods that are already being produced. Bohm<sup>57</sup> suggests that people be posed questions where the payment rule is deliberately specified as yet-to-be-determined. In this manner, he expects to cancel the incentives to over- or underrepresent true feelings, because people will not be able to select a strategy for a misrepresentation with the aim of making themselves better off than telling the truth.

Bohm<sup>58</sup> conducted an experiment to see how sensitive willingness-to-pay responses were to question wording and to analyze whether strategic behavior seemed present. The sample does not purport to be fully representative (only 211 of 605 randomly selected residents of Stockholm agreed to participate),

---

53. Lindahl, *Some Controversial Questions in the Theory of Taxation*, reprinted in *CLASSICS IN THE THEORY OF PUBLIC FINANCE* 214 (R. Musgrave & A. Peacock eds. 1958).

54. J. ACTON, *supra* note 18.

55. Bohm, *supra* note 33.

56. Dreze & Poussin, *A Tatonnement Process for Public Goods*, 38 *REV. ECON. STUD.* 133 (1971).

57. Bohm, *supra* note 33.

58. P. Bohm, *Estimating Demand for Public Goods: An Experiment* (n.d.) (unpublished paper on file at Department of Economics, University of Stockholm).

but the experimental design is intriguing and to the point. He paid the volunteers Kr.50 (ten dollars) for a one hour "interview" about television programs. When the respondents came to the studio, they were told the interview was delayed and they were put in a room with television screens and given an opportunity to watch a comedy show with two very popular comedians. They were given the impression that several other respondents were in similar rooms around the building and that the program would be shown only if the aggregate willingness to pay exceeded the cost associated (Kr.500). The different respondents were randomly given different instructions about what the decision rule for actual showing would be.<sup>59</sup> If people were behaving strategically, some instructions should cause significantly higher responses than other instructions. Bohm's empirical results show no statistically significant difference (at five per cent) in the responses from one question form to another.

At the moment, one can conclude that, although strategic misrepresentation may exist in principle in the willingness-to-pay context, it has not been demonstrated to be a significant empirical factor. At the pragmatic level, it is relatively unlikely to be a serious problem with preliminary efforts to assess people's values, because people are not accustomed to having their tax bill react to such statements of value.<sup>60</sup>

Many of these potential problems in implementing a willingness-to-pay measure will be clarified only with additional empirical evidence. For instance, the estimates of the true variance of responses in society and the mean value

---

59. For instance, "you pay your actual maximum willingness to pay," or "you pay some fraction," or "you pay a proportion yet-to-be-determined," and so forth.

60. Other means besides a willingness-to-pay survey can be used to elicit the explicit values of individuals, but none of them answers the operational question of evaluation: How much should be spent on programs that change people's chances of death or disability? The exception to this assertion is a scaling technique that employs von Neumann-Morgenstern lotteries to determine a utility function. It has been demonstrated that this is formally equivalent to a willingness-to-pay approach. C. Neu, *The Use of Individual Preferences in the Public Valuation of Life and Health* (1975) (unpublished dissertation in Department of Economics, Harvard University). The remaining techniques cannot provide the operationally needed answer. For instance, a variety of psychometric scaling devices could be employed to measure people's attitudes toward attributes of program impact (say, death or disability), or their attitudes toward programs (say, heart attack ambulance or anti-hypertension programs). The results of such a scaling, however, do not answer the fundamental question of evaluation: Should scarce resources be committed? Suppose I know that Program *A* scores 8 and Program *B* scores 4 on a 10 point scale where 0 is very bad and 10 is very good. I still do not know whether or not to undertake either program. Suppose one includes information about program cost and defines the status quo as 5 on the scale, one would still not know if either program should be undertaken. Furthermore, even if such a scaling produced an indication that a program should or should not be undertaken, the results are of limited applicability because one knows only the valuation of a few programs rather than having a procedure that can be generalized. Another approach would be to ask people if they would like to see more, less, or the same amount spent on a given public program. If one then asked how much more should be spent, and specified the person's share of the cost, one would have a result equivalent to willingness-to-pay results and would answer the question of evaluation. Furthermore, if one asks enough questions, this iteration will produce a majority rule situation, which has significant appeal as a public decision making criterion.



of the responses can only be judged by conducting surveys on representative populations of respondents. Similarly, the reproductibility and stability of responses over time can be measured, but have not yet been explored empirically. Some of the more basic concerns about the validity of the responses and the internal consistency of a given person's responses are more difficult to resolve. Analysts have crude measures of what "internal consistency" means, but to demonstrate rigorously its existence (or nonexistence), hard thinking is needed. An interactive process of both conceptual development and refined empirical evidence seems to be the most suitable strategy for furthering understanding in both areas. Furthermore, if done with some foreplanning, analysts can also provide useful interim survey results that can serve as one measure of social impact valuation for current evaluation efforts.

#### CONCLUSION

There are important conceptual and empirical differences among the approaches to evaluation reviewed here. The choice of method is important and can change the ranking and value of health or safety programs significantly. The selection of a particular method involves tradeoffs between ease of application and conceptual soundness. The livelihood saving approach is easy to apply (and has been used frequently in the past), but it has a number of drawbacks when its implications are examined in detail. An approach based on individual preferences (operationally, what people are willing to pay) meets the drawbacks of the livelihood approach and is conceptually most satisfactory. Preliminary evidence suggests that it is feasible to ask for explicit statements and that meaningful answers result, but a number of problems may arise in implementation on a large scale. There has been very little empirical experience with measuring implicit value or with conducting surveys of people's willingness to pay for public programs. In the revealed preference approaches analysts may not observe a representative group of people, and it may be difficult to know with certainty that observed behavioral differences should be attributed only to differences in level of risk. Correspondingly, analysts do not know what the stability of survey responses is over time or what the sample variance is likely to be. Furthermore, the validity and internal consistency of these responses is not yet established. It is difficult to specify rigorous tests of the external validity of these sorts of questions, but an interactive development of the conceptual underpinnings and empirical evidence provides promise of sharpening our understanding.

For many actual evaluations, both the livelihood saving approach (with its known drawbacks) and an imperfect, crudely measured, willingness-to-pay methodology are clearly superior to no formal analysis. First, the analysis is frequently an order-of-magnitude evaluation. Under these circumstances, the drawbacks or questions of either approach are second-order magnitudes and

do not affect the conclusion whether or not to undertake the program. Second, employing both criteria to see if they yield the same conclusion can reinforce one's confidence in the robustness of the decision. Third, in the range of expected effectiveness for many realistic programs, the approaches frequently lead to reasonably close measures of value.<sup>61</sup>

When given a choice between livelihood saving or willingness to pay as a basis for evaluating social impact, a strong case can be made for the conceptual superiority of willingness to pay. The livelihood measure does not bear any necessary relationship to what people want in the way of public programs. If it is decided to fund programs by this criterion, it could, in general, raise adequate revenues by taxing those whose livelihood is extended.<sup>62</sup> However, this criterion does not guarantee that society or any individual is made better off by adopting the program.

An individual preference approach (based on willingness to pay) does provide an assurance that society is made better off in some sense by the programs that pass the criterion. By approving only programs such that people are willing to pay, in the aggregate, more than the programs cost, we can make a strong case that society as a whole gains. It is clear that in general the program will be funded in a manner such that some people gain and some lose with a particular implementation. Nevertheless, since the aggregate willingness to pay exceeds the cost, it would be possible to spread the costs such that no one was made worse off by the program. That is, this criterion identifies potential Pareto superior moves for society. Every member can be at least as well off as he was without the program, and at least one person is better off.

Although this article began with the objective of identifying means of placing a value on reductions in probability of death or disability, it may not be possible (or desirable) to have a unique value that can be used in several different contexts. Instead, it may turn out that preferences are such that there is one value for a change in probability for cancer death, another value for a change in probability of heart attack death, and yet a third value for change in probability of accidental death—even for similar persons and identical starting risks and reduction in risks. Given the diversity of values now implicit in public decision making, such a finding would not be unexpected. Furthermore, analysts like Zeckhauser<sup>63</sup> argue that the process by which public decisions are made may be at least as important as the actual numerical values used. An appropriate strategy for the decision maker charged with evaluating

---

61. For instance, in J. ACTON, *supra* note 18, the conclusions as to net benefit of five interventions for out-of-hospital heart attacks were very similar under both methods of evaluation.

62. That is, if the government were to tax away an amount up to the entire future earnings of individuals whose lives were saved, then it would cover the costs of such programs. In the absence of indentured servitude, even this situation may not always be realized.

63. Zeckhauser, *supra* note 1.

lifesaving programs before additional methodological and empirical research takes place may be to employ more than one of the techniques discussed. When the different approaches yield similar conclusions, he can gain confidence from the fact that his evaluation does not seem to be sensitive to the values employed. When they yield sharply different conclusions, he can probe his own preferences or seek additional evidence about the willingness of the target population to pay.